DOCUMENT RESUME

ED 352 392 TM 019 296

AUTHOR Hewett, Stephenie M.; Brady, Mary Ella
TITLE An Analysis of Statewide NRT Programs.
PUB DATE Mar 92
NOTE 35p.; A previous version of this paper was presented at the Annual Meeting of the Eastern Educational Research Association (15th, Hilton Head, SC, March 5-9, 1992).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Achievement Tests; Educational History; Educational Policy; Elementary Secondary Education; National Surveys; *Norm Referenced Tests; Program Evaluation; *Standardized Tests; *State Programs; State Surveys; Statistical Data; *Testing Programs; Test Norms; *Test Use

ABSTRACT

State program practices in norm-referenced testing (NRT) are described as found in the 1990-91 school year and prior years to provide information on years of administration of the same test and the age of the norms. All 50 State Departments of Education were surveyed in 2 stages. States without NRT were asked to described their testing policies; and those with NRT were asked test name, editions, grades tested, and first year of administration. Only two states provided charts of the tests administered, and information on the programs before the 1980s was often incomplete. In 1991, 29 states had statewide NRT programs, most of which tested target populations in the spring. The oldest mandatory program began in 1955, although optional statewide testing existed in many states before mandatory programs were established. Four of five states gave one of three tests (Comprehensive Tests of Basic Skills, Iowa Tests of Basic Skills, and Stanford Achievement Tests). Of the 29 states, 23 gave the most current edition to the researchers and the others used the immediate prior edition. A majority (57%) used NRTs normed 3 years previously, but just over one-third used norms 6 years old. Across all states, the average number of years of administration of the same test was 3.6 years. It is recommended that NRTs be renormed in 3-year cycles, and that tests be adopted for 3- or 4-year cycles. Three tables contain survey data. Appendix A lists the tests used by state. Appendix B lists the grades and subtests by state. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED352392

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

STEPHENIE M. HEWETT

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

AN ANALYSIS OF STATEWIDE NRT PROGRAMS

Stephenie M. Hewett, Ed.D.
The Citadel
Charleston, South Carolina
and
Mary Ella Brady, Ph.D.
Charleston County School District
Charleston, South Carolina

A previous version of this paper was presented at the Eastern Educational Research Association annual meeting in Hilton Head, S.C., March 7, 1992. The research reported here was supported in part by a grant from The Citadel. The opinions expressed here are solely those of the authors; no endorsement by The Citadel or the Charleston County School District is implied.

TM019296

BEST COPY AVAILABLE

AN ANALYSIS OF STATEWIDE NRT PROGRAMS

Educational leaders have long recognized that to gain public support for education they must be able to demonstrate successful learning in the schools. Standardized achievement tests were first developed more than 50 years ago so that consistent measures of achievement in schools could be used. The availability of these tests led to districtwide testing programs and, later, to statewide testing programs.

Both district and statewide programs can use only norm-referenced tests (NRT) or criterion-referenced tests (CRT), or combinations of the two at different grades (Linn and Hambleton, 1991). Recent references on testing provide sketchy information on current statewide program practices. The attention drawn by Cannell (1989) to state NRT results, including his accusation that educators were cheating on standardized tests because every state reported average achievement above the national average, suggests a need to examine current statewide NRT practices with special attention to the age of norms of NRTs used in state programs.

Objectives

The goal of the study was to describe statewide program practices

in norm-referenced testing as they existed in 1991 (1990-1991) and in prior years. The primary objective was to identify practices in place for the 1991 school year in NRT statewide testing programs as far as tests used, age of norms, grade levels, subject areas, and number of years tests had been in use. The secondary objective was to identify NRT tests used since the start of statewide programs and their editions, norm years, and years administered. Thus, the intent was to provide information on two factors that are important considerations in designing testing programs --- years of administration of the same test and age of norms.

Background

With the recent publication of Understanding Achievement Tests (Rudner, Conoley, and Blake, 1989), school district staff were provided with a source of information on standardized achievement tests and testing procedures. This guide is a valuable reference, as it includes sections on misuses of tests, test preparation practices, and related topics, as well as charts describing major tests and publisher lists. However, specific information on which standardized achievement tests states actually administer and state administration practices is not included.

For states using NRTs, the test used, the time of year administered, the grade levels involved, and the subject areas tested vary to such an extent that each state's program is virtually unique, according to CCSSO (1989). Because there are only five NRTs that can be considered major norm-referenced achievement tests, some overlap in tests given must exist across states with NRT testing programs. Also, a majority of states test in the spring, leaving little variability for time of year. Very few states can be expected to test all twelve grades, given the expense involved, but some overlap is expected when levels are considered, e.g., primary, elementary, middle, and high schools. It is clear that similarities at a general level must exist in state programs.

Four recent sources provide information on statewide NRT programs. Cannell (1989) provided lists of states giving specific NRT tests, as did Shepard (1989), who also gave the first year of administration by state and sporadic data on forms, grades, and subjects administered. While these two sources do contain information about state testing practices, the discrepancies between them on tests given, and the lack of indepth information reveal the need for a more up-to-date, detailed survey of state testing

practices.

Such a survey was conducted by the Southern Regional Educational Board (Korcheck, 1988) for states in the South, while a national telephone survey was conducted by Palavin Associates. The former addresses only Southern states, and the latter is not part of the traditional literature. Information on the frequency and type of changes in norm-referenced achievement tests in states is also lacking.

Cannell (1989) concluded that there was nationwide cheating on NRT tests because, according to him, each state giving NRTs reported scores above the national average. However, other investigators (Linn et al, 1991; Wiser and Lenke, 1987) have found evidence that the longer the same form and edition of a test is used, the greater the improvement. This was labeled the "creeping upwards" effect, with age of the test norms being one reason for the upwards movement and test familiarity, a second. If, for example, in 1991 the average years of administration of the same test for states with NRT programs was five, there would clearly have been an opportunity for the "creeping upwards" effect to operate. Age of norms of tests in use is directly related to years of administration of the same test and also warrants

investigation.

Methods

All 50 state departments of education were surveyed in two stages. In the initial stage, information was gathered on whether or not they currently had an NRT statewide testing program, i.e., whether or not they mandated use of a single norm-referenced test to populations or samples at least once every three years. States without an NRT program were asked to describe their statewide testing policy, if any. States with an NRT program were asked to provide the names of NRT test(s) used in 1991, the editions, grades tested, and first year of administration. Similar information was requested of all previous tests used since the start of the statewide testing program. In the second stage, telephone calls to more than 30 states were used to clarify state practices and to request 1991 state reports.

The three sources of information were personnel in student assessment units of state departments of education, publishing company personnel, and the documents sent. Documents ranged from one page tables to actual test summary reports produced by publishers to annual reports. Inconsistencies between sources were

resolved as much as possible by phone calls.

Results

The first finding of the study was that gathering historical information just on the bare facts of statewide testing programs was more difficult than anticipated. Only two states (South Carolina and Virginia) provided charts of the tests administered since the NRT program began. One state could not provide the year the program began, except to say "Many moons ago". Information on programs in the 1980's was almost complete, while information prior to then was often incomplete. For that reason grade and subtest data were collected only for 1991 NRT statewide testing programs.

Appendix A gives basic information on individual statewide NRT programs for 1991 and as far back as data were available. (Missing data are indicated by '?'.) Information is provided for year began, test, norm year, and years of administration. In 1991, 29 states had statewide NRT programs. Most of the programs tested complete populations in the target grades every spring. A few states tested samples of students, while one state, Missouri, tested samples of students and samples of items. Several states tested in the fall, e.g., Washington, and two (Colorado and Utah) tested every third year.

Almost all of the 1991 statewide NRT programs began within the last twenty years. The oldest program was begun in Virginia in 1955, while the newest one began in North Dakota in 1990. Almost half of the programs began in the 1980's. The fact that states were asked when they began **mandated** use of a single norm-referenced test statewide probably accounts for the relative newness of the programs. Optional statewide testing existed previously in many states, while mandatory testing in districts existed long before state programs were set up.

In 1971, six of the 29 states that had NRT programs in 1991 were already operating statewide programs. Complete information was obtained only for Georgia, which administered the ITBS/TAP (5 or 6) with 1970 norms for the first time in 1971. By 1976, nine states had statewide programs; six of these provided complete data, as follows:

Alabama: CAT-A,B (1970 norms), third administration;
Georgia: ITBS-5,6 (1970 norms), sixth administration;
Hawaii: SAT-6 (1972 norms), first administration;
Maryland: ITBS-5,6 (1970 norms), fourth administration;
New Mexico: CTBS-S (1973 norms), second administration; and
Utah: CTBS-S (1973 norms), first administration in 1975 of a three year cycle.

In these six states, four different tests were administered. The number of years the same test had been administered ranged from one

to six years with a mean of 2.8. Age of norms ranged from three to six years, with a mean of 4.7.

By 1981, 16 of the 29 states were operating statewide NRT programs. Fifteen provided complete data as follows:

Alabama: CAT-C (1977 norms), third administration;
Arizona: CAT-C (1977 norms), first administration;
Arkansas: SRA-1 (1978 norms), second administration;
Delaware: CAT-C (1977 norms), third administration;
Georgia: ITBS/TAP-7,8 (1978 norms), fourth administration;
Hawaii: Stanford-6 (1972 norms), sixth administration;
Kentucky: CTBS-S (1973 norms), third administration;
Maryland: CAT-C (1977 norms), first administration;
New Mexico: CTBS-S (1973 norms), seventh administration;
North Carolina: CAT-C (1977 norms), fourth administration;
Rhode Island: ITBS-7,8 (1978 norms), first administration;
South Carolina: CTBS-S (1973 norms), fourth administration;
Utah: CTBS-S (1973 norms), third administration;
Washington: CAT-C (1977 norms), second administration; and
West Virginia: CTBS-S (1973 norms), fifth administration.

The CAT and CTBS were by far the most popular tests in 1981, as they were given in 11 of 15 states. There was wide variability in the number of times NRTs had been administered, i. e., from one to seven years with a mean of 3.3. Norm age ranged from three to nine years with a mean of 5.5.

In 1986, 23 of the 29 states had statewide NRT programs in operation. Twenty-two of the states supplied complete data as follows:

Alabama: Stanford-7 (1982 norms), second administration;
Arizona: Stanford-7 (1982 norms), second administration;
Arkansas: MAT-6 (1985 norms), first administration;
Delaware: CTBS-U, V (1981 norms), third administration;
Georgia: ITBS/TAP-G (1985 norms), first administration;
Hawaii: Stanford-7 (1982 norms), second administration;
Idaho: ITBS/TAP-G (1985 norms), second administration;
Kentucky: CTBS-U (1981 norms), fifth administration;
Maryland: CAT-C (1977 norms), sixth administration;
Mississippi: Stanford-7 (1982 norms), second administration;
New Hampshire: CAT-E (1985 norms), first administration;
New Mexico: CTBS-U (1981 norms), fifth administration;
North Carolina: CAT-E (1985 norms), first administration;
Oklahoma: MAT-6 (1985 norms), first administration;
Rhode Island: MAT-6 (1985 norms), first administration;
South Carolina: CTBS-U (1981 norms), fourth administration;
South Dakota: Stanford-7 (1982 norms), second administration;
Tennessee: Stanford-7 (1982 norms), second administration;
Utah: CTBS-U (1981 norms), first administration (Utah tested every three years and last tested in 1984.);
Virginia: SRA-1 (1978 norms), fifth administration;
Washington: MAT-6 (1985 norms), first administration; and
West Virginia: CTBS-U (1981 norms), second administration.

In 1986, just over half of the states with NRT programs administered the Stanford-7 or the CTBS-U, V. The number of years that the same test had been administered ranged from 1 to 6 with the mean number of years of administration being 2.4 years. Norm age ranged from 1 to 9 years with the mean norm age being 3.5.

Table 1 shows the NRTs used in the 29 states with statewide programs in 1991. Four out of five states gave one of three tests. Eight gave the Comprehensive Test of Basic Skills, eight, the Iowa

Table 1

NRTs in Statewide Testing Programs in 1991

Test	Number	States
California Achievement Test	(3)	
CAT-E	2	New Hampshire, North Carolina
CAT-F	1	Louisiana
Comprehensive Test Of Basic Skills	(8)	
CTBS-U	1	West Virginia
CTBS-4	7	Indiana, Kentucky, Maryland, Nevada, New Mexico, North Dakota, Tennessee
Iowa Test of Basic Skills	(8)	
ITBS-G	5	Alaska, Colorado[1] , Idaho, Missouri, Virginia
ITBS-J	3	Arizona, Georgia, Oklahoma
Metropolitan Achievement Test	(3)	
MAT-6	3	Arkansas, Rhode Island, Washington
Stanford Achievement Test	(7)	
Stanford-7	1	Hawaii
Stanford-8	6	Alabama, Delaware, Mississippi South Carolina, South Dakota, Utah
Total	29	

[1] Colorado tests every third year. This was the test given in the Fall of 1988.

Test of Basic Skills, and seven, the Stanford Achievement Test. Twenty-three of the 29 states were giving the most current edition available, while six were giving the immediate prior edition.

A majority (57%) of states used NRTs in 1991 that had been normed three years previously, as shown in Table 2. Just over one-third of the states were using NRTs with norms that were six years old, while less than 10% used tests with norms that were nine or ten years old. The average age of NRT norms in 1991 was 4.5 years. This compares to means of 4.7 in 1976, 5.5 in 1981, and 3.5 in 1986. Almost three-fourths of these NRT states had been administering the same test less than six years. Across all states, the average years of administration of the same test was 3.6. This compares to 2.8 in 1976, 3.3 in 1981, and 2.4 in 1986.

As expected, the fewer the number of years of administration of the same test, the younger the norms. Thus, of the twelve states administering an NRT for the first or second time in 1991, norm age was three for each. Three states-- Hawaii, Idaho, and West Virginia-- had each administered the same NRT the longest, i. e., seven years. Norms were nine, six, and ten years old, respectively.

In addition to norm age, grades and subjects tested are also

Table 2

Age of Norms and Number of Years Same NRT Administered in Statewide NRT Programs as of 1991[1]

Age of Norms		Number of Years Same NRT Administered	
Age	Number of States (%)	Years	Number of States (%)
1	0 (0%)	1	3 (11%)
2	0 (0%)	2	9 (32%)
3	16 (57%)	3	5 (18%)
4	0 (0%)	4	2 (7%)
5	0 (0%)	5	1 (4%)
6	10 (36%)	6	5 (18%)
7	0 (0%)	7	3 (11%)
8	0 (0%)	8	0 (0%)
9	1 (4%)	9	0 (0%)
10	1 (4%)	10	0 (0%)
11+	0 (0%)	11+	0 (0%)

[1] Colorado, which tests every three years and last tested in the Fall of 1988, is excluded.

important characteristics of statewide NRT programs. Data are provided in Appendix B and summarized in Table 3. Subtests were coded if they were reported, even if they were part of a composite score. For example, spelling was often part of language but both spelling and language were coded if they were both reported.

Each of the 29 NRT states mandated testing in 1991 in more than one grade. One state--Alabama-- tested only two grades, and none tested just one grade. Almost half of the states tested in three grades. States testing the most number of grades were Arizona (all twelve), Delaware (grades 1-8 and 11), and Missouri (grades 2-10). As shown in Table 3, the first, second, and twelfth grades were tested in the fewest states. All the states tested one or more of the elementary and middle school grades, while 8 out of 10 states tested one or more high school grades. The single grade tested by most of the states was grade 8 (69%). This is not surprising given that grade 8 is the last year prior to high school.

As with grades, each of the 29 NRT states mandated testing in more than one subtest area. Just over one-fourth of the states administered complete batteries. All states administered reading (or reading comprehension), mathematics, and language, while 8 out of 10

Table 3

Grades Tested and Subtests Given in
Statewide NRT Programs in 1991[1]

Grade	Number of States(%)	Subtest[2]	Number of States (%)
1	4 (14%)	Listening	8 (28%)
2	7 (24%)	Word Analysis	12 (41%)
3	15 (52%)	Vocabulary	20 (69%)
4	16 (55%)	Reading Comprehension	22 (76%)
5	10 (34%)	Reading	27 (93%)
6	15 (52%)	Math	29 (100%)
7	10 (34%)	Language	29 (100%)
8	20 (69%)	Social Studies	20 (69%)
9	9 (31%)	Science	20 (69%)
10	10 (34%)	Spelling	21 (72%)
11	11 (38%)	Study Skills	15 (52%)
12	1 (3%)	Basic Battery	23 (79%)
		Complete Battery	8 (28%)

1 Colorado, which tests every three years and last tested in the Fall of 1988, is included.

2 Subtests are counted when they were administered in at least one grade and were reported.

reported Basic Battery scores, which are a composite of these three subtests. Approximately 7 out of 10 states administered social studies, science, and spelling. It must be noted that subtests administered are controlled by the test selected, e. g., only Stanford includes listening beyond the primary grades. Also, Table 3 counts a subtest only one time, no matter how many grades in which it was administered. This masks the fact that the non-core areas such as study skills tend only to be given at the higher grades. Appendix B, however, does show individual subtests by grade.

Discussion

The difficulties faced in the current study in getting historical information on statewide NRT programs may not be a problem in the future if states submit annual reports to ERIC, as many do now. For the present study, the focus had to shift to a general analysis of recent statewide NRT practices because detailed historical data could not be obtained.

One unanticipated finding of the study was the relative newness of most statewide NRT programs. Obviously, if all states had been surveyed about the existence of mandated NRT programs prior to their 1991 mandated CRT programs, the results would have been different.

States such as California, Michigan, and New York had extensive statewide CRT programs in 1991 and prior to CRT implementation had NRT programs. Another reason for relative newness of programs probably is the definition of statewide NRT program used in the current study: mandated statewide administration as opposed to voluntary administration. A follow-up study addressing both of these issues is warranted.

The data suggest that norm-referenced statewide testing has dramatically increased in the last twenty years, i. e., from 6 to 29 states. It is doubtful that this is the case, however, because historical data were collected only on states with NRT programs in 1991. If there were 10 NRT states in 1971 that subsequently became CRT states, the increase would be less than dramatic. In any event, it is doubtful that dramatic increases will occur in the future because almost all of the remaining states have mandated CRT statewide programs, some of which were set up to replace NRT programs. Also, the expansion of NAEP to the state level and the proposed national test of educational goals will probably slow the adoption of NRT statewide programs. However, marked decreases in states with mandated NRT statewide programs is also unlikely.

The general public, including parents, will always want an answer to the question, "How are we doing compared to the nation?" The national percentile ranks of NRTs provide a relatively understandable answer to this question. As many people do not understand the concept of sampling, the need for population testing arises. Also, if NRT results are, somewhat incorrectly, going to be used to make instructional decisions about every individual child, population testing is again needed. The logistical difficulties of doing sophisticated statewide sampling may be a third reason few states use sampling.

There were fewer testing companies in 1991 than in 1971, which may explain why 4 out of 5 NRT states used one of three tests in 1991. The norming cycles of these tests have hardly changed over the years; thus, the age of norms used in statewide NRT programs in 1991 was only 0.2 years younger than the norms of 1976. States, however, are in a sense "prisoners" of the publishers of NRTs and must buy what the publishers make available. For the three most popular tests in statewide NRT programs in 1991, a sharp drop occurred in the number of years between renorming for the edition current as of 1991 and the immediate previous edition for Stanford and ITBS but not for

CTBS. The Stanford-6 was published 10 years before the Stanford-7 was introduced, but Stanford-8 came out 6 years after Stanford-7. The drop for ITBS was from 7 years between ITBS-7, 8 and ITBS- G, but three years between ITBS-G and ITBS-J. Additional evidence that publishers are norming more frequently is the fact that Stanford-8 was renormed in 1991. More frequent norming should result in more frequent test changes in state programs in the future and, thus, more up-to-date norms.

The findings of this study concerning grades and subtests administered in statewide NRT programs in 1991 were not surprising. NRT testing at the primary grades has not been well supported in the past and is receiving even less support now. Testing the high school grades is not very valuable because the weak link between NRT subtests and high school courses means that little useful instructional information is given by an NRT. Also receipt of a high school diploma is viewed as a more desired goal than a high percentile on an NRT. Most states test elementary and middle grades that are "gateposts", e.g., grades 4 and 8, with the basic skill areas related to literacy and numeracy, areas that have received the most media focus.

In conclusion, it is hoped that states and districts adopting NRTs will find the information here useful. While the reasons behind state practices are not given, knowing what other states are doing should assist the decision making process. It is also hoped that the number of years of administration of the same test, along with age of norms, will decrease in the future. Given the "creeping upwards" effect, administering the same test for seven years is not educationally sound. Nor is the use of 10 year old norms, especially during a time period when achievement, as measured by NRTs, has been increasing. It is recommended that all NRTs be renormed in three year cycles, and that states or districts adopt NRTs for three or four year cycles.

References

Cannell. J.J. (1989). How Public Educators Cheat on Standardized Achievement Tests. Albuquerque, NM: Friends for Education.

CCSSO (1989). State Education Indicators: 1989. Washington, D.C.: Council of Chief State School Officers.

Hymes, D.L.; Chafin, A.E ; and Gonder, P. (1991). The Changing Face of Testing and Assessment. American Association of School Administrators. Arlington, Virginia.

Korcheck, S.A. (1988). Measuring Student Learning: Statewide Student Assessment Programs In the SRB States. Southern Regional Educational Board, Atlanta, Georgia.

Linn, R.L., and Hambleton, R.K. (1991). Customized Tests and Customized Norms. Boulder, Colorado: University of Colorado/CRESST.

Linn, R.L.; Graue, M.E.; and Sanders, N.M. (1990). Comparing state and district results to national norms: The validity of the claims that "Everyone is above average". Educational Measurement: Issues and Practice, 9, 5-14.

Rudner, L.M.; Conoley, J.C.; and Plake, B.S. (1989). Understanding Achievement Tests: A Guide for School Administrators. Washington, D.C.: The ERIC Clearinghouse on Tests, Measurement, and Evaluation, AIR.

Shepard, L.A. (1989). Inflated test score gains: Is it old norms or teaching the test? Los Angeles, CA: UCLA Center for Research on Evaluation, Standards, and Student Testing.

Wiser, B. and Lenke, J.M. (1987). The stability of achievement test norms over time. Paper presented at the annual meeting of the National Council on Measurement in Education, Washington, D.C.

APPENDIX A

TESTS GIVEN IN STATE NRT TESTING PROGRAMS AS OF 1991

STATE	YEAR BEGAN	TEST	NORM YEAR	YEARS GIVEN
Alabama	1958	Stanford-8	1988	90-91
		Stanford-7	1982	85-89
		CAT-C	1977	79-84
		CAT-A or B	1970	74-78
		?	?	58-73
Alaska	1989	ITBS-G	1985	89 & 90 or 91[1]
Arizona	1981	ITBS/TAP-J	1988	89-91
		ITBS/TAP-H	1985	88
		Tests of Academic Skills & Knowledges[2]	?	85-87
		Stanford-7[2]	1982	85-87
		CAT-C	1977	81-84
Arkansas	1980	MAT-6	1985	86-91
		SRA-1	1978	80-85
California	*NA*			
Colorado	1988	ITBS/TAP-G	1988	89[3]
Connecticut	*NA*			
Delaware	1979	Stanford-8	1988	89-91
		CTBS-U,V	1981	84-88
		CAT-C	1977	79-83
Florida	*NA*			
Georgia	1971	ITBS/TAP-J	1988	90-91
		ITBS/TAP-G	1985	86-89
		ITBS/TAP-7 or 8	1978	78-85
		ITBS/TAP-5 or 6	1970	71-77

1 Alaska had administered the ITBS-G twice as of 1991. For the second administration, some districts tested in Spring 1990 and some in Fall 1990.

2 Both tests were given from 85 to 87; further information was not available.

3 Colorado tests every third year.

APPENDIX A (CONTINUED)

TESTS GIVEN IN STATE NRT TESTING PROGRAMS AS OF 1991

STATE	YEAR BEGAN	TEST	NORM YEAR	YEARS GIVEN
Hawaii	*Unknown*	Stanford-7	1982	85-91
		Stanford-6	1972	76-84
		STEP-?	?	?-75
Idaho	1985	ITBS/TAP-G	1985	85-91
Illinois	*NA*			
Indiana	1987	CTBS-4	1988	91
		CAT-E	1985	87-90
Iowa	*NA*			
Kansas	*NA*			
Kentucky	1979	CTBS-4	1988	89-91
		CTBS-U	1981	82-88
		CTBS-S	1973	79-81
Louisiana	1988	CAT-F	1985	88-91
Maine	*NA*			
Maryland	c1973	CTBS-4	1988	91
		CAT-C	1977	81-90
		ITBS/TAP-5 or 6	1970	73-80
Massachusetts	*NA*			
Michigan	*NA*			
Minnesota	*NA*			
Mississippi	1985	Stanford-8	1988	90-91
		Stanford-7	1982	85-89
Missouri	1987	ITBS/TAP-G	1985	87-91

APPENDIX A (CONTINUED)

TESTS GIVEN IN STATE NRT TESTING PROGRAMS AS OF 1991

STATE	YEAR BEGAN	TEST	NORM YEAR	YEARS GIVEN
Montana	*NA*			
Nebraska	*NA*			
Nevada	1980	CTBS-4	1988	90-91
		Stanford-7 (gr. 3 & 6)	1982	?-89
		SRA (gr. 9)	1985	?-89
		?	?	80- ?
New Hampshire	1986	CAT-E	1985	86-91
New Jersey	*NA*			
New Mexico	1972	CTBS-4	1988	89-91
		CTBS-U	1981	82-88
		CTBS-S	1973	75-81[1]
		CTBS-Q	1968	72-74
New York	*NA*			
North Carolina	1978	CAT-E	1985	86-91
		CAT-C	1977	78-85
North Dakota	1990	CTBS-4	1988	90-91
Ohio	*NA*			
Oklahoma	1986	ITBS/TAP-J	1988	90-91
		MAT-6	1985	86-89
Oregon	*NA*			
Pennsylvania	*NA*			
Rhode Island	1969	MAT-6	1985	86-91
		ITBS-7 or 8	1978	81-85
		?	?	?-80
		?	?	69- ?

[1] The CTBS-U was introduced in New Mexico between the 1981 and 1982 school years.

APPENDIX A (CONTINUED)

TESTS GIVEN IN STATE NRT TESTING PROGRAMS AS OF 1991

STATE	YEAR BEGAN	TEST	NORM YEAR	YEARS GIVEN
South Carolina	1978	Stanford-8	1988	90-91
		CTBS-U	1981	83-89
		CTBS-S	1973	78-82
South Dakota	1985	Stanford-8	1988	89-91
		Stanford-7	1982	85-88
Tennessee	1985	CTBS-4	1988	90-91
		Stanford-7	1982	85-89
Texas	*NA*			
Utah	1975	Stanford-8	1988	91[1]
		CTBS-4	1988	90
		CTBS-U	1981	84 & 87
		CTBS-S	1973	75, 78, 81
Vermont	*NA*			
Virginia	1955	ITBS/TAP-G	1985	88-91
		SRA-1	1978	82-87
		SRA-E, M	?	73-81
		SRA-C	?	66-72
		Misc.	?	55-65
Washington	1977	MAT-6	1985	86-91
		CAT-C	1977	80-85
		CTBS-S	1973	77-79
West Virginia	1962	CTBS-U	1981	85-91
		CTBS-S	1973	77-84
		EDS-?	?	70-76
		?	?	62-69
Wisconsin	*NA*			
Wyoming	*NA*			

[1] Utah tested every third year in the spring through 1990. The 1991 testing was in the fall of 1990.

APPENDIX B

GRADES AND SUBTESTS IN STATEWIDE NRT PROGRAMS IN 1991

STATE	TEST	GRADES	SUBTESTS[1]												
			Li	WA	V	RC	R	M	L	SoS	Sc	Sp	StS	BB	CB
Alabama	Stanford-8	4 & 8	Y	N	N	N	Y	Y	Y	Y	Y	N	N	Y	Y
Alaska	ITBS-G	4, 6, & 8[2]	N	N	N	N	Y	Y	Y	N	N	Y	N	N	N
Arizona	ITBS-J	1 & 2	Y	Y	Y	Y	N	Y	Y	N	N	Y	Y	Y	N
	ITBS-J	3-8	N	N	Y	Y	N	Y	Y	N	N	Y	Y	Y	N
	TAP-J	9-12	N	N	N	Y	N	Y	Y	N	N	N	Y	Y	N
Arkansas	MAT-6	4, 7, & 10	N	N	N	N	Y	Y	Y	Y	Y	N	N	Y	N
Colorado	ITBS-G	4 & 7	N	N	Y	Y	N	Y	Y	N	N	Y	Y	Y	N
	TAP-G	10	N	N	N	N	Y	Y	Y	N	N	N	Y	Y	N
Delaware	Stanford-8	1	Y	Y	N	Y	Y	Y	Y	N	N	Y	N	Y	N
		2	Y	Y	Y	Y	Y	Y	Y	N	N	Y	N	Y	N
		3	Y	Y	Y	Y	Y	Y	Y	N	N	Y	Y	Y	N
		4-8	Y	N	Y	Y	Y	Y	Y	N	N	Y	Y	Y	N
		11	N	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Georgia	ITBS-J	2	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	N
	ITBS-J	4 & 7	N	N	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	N
	TAP-J	9	N	N	N	N	Y	Y	Y	Y	Y	N	Y	Y	N
Hawaii[3]	Stanford-7	3 & 6	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y
		8	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y
		10	N	N	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y

[1] Li = Listening, WA = Word Analysis, V = Vocabulary, RC = Reading Comprehension, R = Reading, M = Mathematics, L = Language, SoS = Social Studies, Sc = Science, Sp = Spelling, StS = Study Skills, BB = Basic Battery, CB = Complete Battery.

[2] Four districts in Alaska tested in Spring, 1990, at grades 3, 5 and 7.

[3] Hawaii did not specify which subtests were given. The data above are estimated from the subtests on the Stanford-7.

APPENDIX B (CONTINUED)

GRADES AND SUBTESTS IN STATEWIDE NRT PROGRAMS IN 1991

STATE	TEST	GRADES	SUBTESTS												
			Li	WA	V	RC	R	M	L	SoS	Sc	Sp	StS	BB	CB
Idaho[1]	ITBS/TAP-G	6 & 8	N	N	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y
		11	N	N	N	N	Y	Y	Y	Y	Y	N	Y	Y	Y
Indiana	CTBS-4	1-3, 6, 8, 9 & 11	N	N	Y	Y	Y	Y	Y	N	N	N	N	Y	N
Kentucky	CTBS-4	3	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	N
		5, 7 & 10	N	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N
Louisiana	CAT-F	4 & 6	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N
		9	N	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N
Maryland[2]	CTBS-4	3, 5, & 8	N	N	N	Y	N	Y	Y	N	N	N	N	N	N
Mississippi	Stanford-8	4, 6, & 8	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Missouri	ITBS-G	2	N	N	N	Y	N	Y	Y	N	N	N	N	N	N
	ITBS-G	3-8	N	N	N	Y	N	Y	Y	Y	Y	Y	N	N	N
	TAP-G	9 & 10	N	N	N	N	Y	Y	Y	Y	Y	N	N	N	N
Nevada	CTBS-4	3 & 6	N	N	N	N	Y	Y	Y	N	N	N	N	N	N
		9	N	N	N	N	Y	Y	N	N	N	N	N	N	N
New Hamphsire	CAT-E	4	N	Y	Y	Y	Y	Y	Y	Y	Y	N	N	Y	N
		8 & 10	N	N	Y	Y	Y	Y	Y	Y	Y	N	N	Y	N

[1] Idaho did not specify which subtests were given. The data above are estimated from the subtests on the ITBS/TAP.

[2] The 1991 Maryland School Performance Program Report shows only these three subtests.

APPENDIX B (CONTINUED)

GRADES AND SUBTESTS IN STATEWIDE NRT PROGRAMS IN 1991

STATE	TEST	GRADES	SUBTESTS												
			Li	WA	V	RC	R	M	L	SoS	Sc	Sp	StS	BB	CB
New Mexico	CTBS-4	3	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	N
		5 & 8	N	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N
North Carolina	CAT-E	1	N	Y	Y	Y	Y	Y	Y	N	N	N	N	N	N
		2 & 3	N	Y	Y	Y	Y	Y	Y	N	N	Y	N	Y	N
		6 & 8	N	N	Y	Y	Y	Y	Y	N	N	Y	N	Y	N
North Dakota	CTBS-4	3	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	N
		6, 8 & 11	N	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N
Oklahoma	ITBS-J	3, 5, & 7	N	N	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y
	TAP-J	9 & 11	N	N	N	N	Y	Y	Y	Y	Y	N	Y	Y	Y
Rhode Island	MAT-6	3	N	Y	Y	Y	Y	Y	Y	N	N	Y	N	Y	N
		6, 8 & 10	N	N	Y	Y	Y	Y	Y	N	N	Y	N	Y	N
South Carolina	Stanford-8	4, 5, 7	Y	N	N	N	Y	Y	Y	Y	Y	Y	Y	Y	Y
		9 & 11	N	N	N	N	Y	Y	Y	N	N	N	N	Y	N
South Dakota	Stanford-8	4 & 8	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
		11	N	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Tennessee	CTBS-4	2-8, & 10	N	N	N	N	Y	Y	Y	Y	Y	N	N	N	N
Utah[1]	Stanford-8	5, 8, & 11	N	N	N	N	Y	Y	Y	Y	Y	N	N	Y	N

[1] As reported in Summary of Initial Results from the Utah Statewide Testing Program 1990.

APPENDIX B (CONTINUED)

GRADES AND SUBTESTS IN STATEWIDE NRT PROGRAMS IN 1991

STATE	TEST	GRADES	SUBTESTS											
			Li	WA	V	RC	R	M	L	SoS	Sc	Sp	StS	BB	CB
Virginia	ITBS-G	4 & 8	N	N	Y	Y	N	Y	Y	Y	Y	Y	Y	N	N
	TAP-G	11	N	N	N	N	Y	Y	Y	Y	Y	N	Y	N	N
Washington	MAT-6	4	N	Y	Y	Y	Y	Y	Y	N	N	Y	N	Y	N
		8 & 10	N	N	Y	Y	Y	Y	Y	N	N	Y	N	Y	N
West Virginia	CTBS-U	3, 6, 9 & 11	N	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N